CHAPTER 1
CHARACTERIZATION OF AMBIENT AIR QUALITY DATA FOR OZONE AND ITS PRECURSORS

1.1   INTRODUCTION

Characterization of ozone episodes is extremely important. A key consideration in devising an effective control strategy is to ensure that the strategy works under a variety of conditions which have been observed to correspond with high measured ozone. The relative importance of ozone/precursors transported into an area is likely to be one important distinguishing characteristic of episodes. Thus, the needs to accurately characterize episodes and identify the potential role of transport are closely related. In a modeling analysis, performed to assess adequacy of proposed control strategies, it is important to have chosen a limited number of episodes which are representative of differing conditions leading to high observed ozone. More generally, a knowledge of meteorological, precursor characteristics and other ambient conditions which correspond with high ozone is useful in helping to formulate policies which are likely to lead to improvements in measured ozone levels. In this chapter, we present examples illustrating the added value of PAMS data in characterizing ozone episodes and identifying features that may be linked to significant pollutant transport.

Both Sections 1.2 and 1.3 focus on how PAMS data can be used to enhance the characterization of local ozone episodic events. The examples presented illustrate the potential uses for and the value-added of both the meteorological and ambient air quality data collected at PAMS sites. The discussions include the extent to which the PAMS data may be useful in assessing the potential role of transport in different episodes.

1.2   EPISODE CHARACTERIZATION USING METEOROLOGICAL MEASUREMENTS[1]

Ozone conducive meteorological conditions such as high insolation, high temperature, high stability (as often reflected by low mixing heights), low winds, and low midday relative humidity have been identified by various researchers in the past (Bruntz et al., 1974; Lamb et al., 1987; Chu, 1987; Chu and Doll, 1991; Cox and Chu, 1993; Robinson, 1952; Hosler, 1961; Ludwig et al., 1977; Pagnotti, 1990). These ozone conducive meteorological conditions can thus be used to characterize the local ozone episodes.

1.2.1 Characterizing Episode "Severity"

EPA has developed relatively simple regression based models to estimate the severity of daily ozone episodes using meteorological data from the National Weather Service. Presently, this methodology considers only surface meteorological variables to characterize episode severity. Since PAMS provides on-site surface as well as upper air meteorological data, the potential for more accurate classification and characterization of distinct episode types (and eventually severity) exists. The routine upper air (rawinsonde) data collected by the National Weather Service (NWS) network are intended to resolve synoptic scale weather systems which have a length scale of about 2000 - 4000 km. Air quality episodes, however, are usually observed in a smaller domain of 1000 km or less in which meso-scale dynamics play a significant role. Thus, the region wide PAMS upper air measurements (in particular, the data collected by RASS and radar profilers) provide valuable information on local mixing height changes, as well as detecting nocturnal jets and topographically induced meso-scale circulations. These, in turn, will help better describe the local characteristics of the episodes and provide inference of possible intra- and/or inter-regional pollutant transport.

Cox and Chu (1993; 1996) have developed a statistical model to predict ozone producing potential using local surface and upper air meteorological data. The model has been applied to minimize meteorological influences on ozone trend analysis, to rank the local severity of ozone episodes, and to select episodic days for modeling and control strategy designs. The PAMS hourly surface and upper meteorological air data (e.g., winds, temperature, humidity, and mixing height data) will certainly increase the power of these models to better define the characteristics of local episodic ozone events as should PAMS air quality data described in Section 1.3. PAMS data, particularly from the RASS and radar profiler sites, will provide better estimates of mixing heights as well as other meteorological measurements aloft. These variables could improve the skill of the statistical model in ranking meteorological ozone forming potential.

1.2.2 Determining Source/receptor Orientations Corresponding to Ozone Conducive Conditions

Chu (1995) has shown that the frequency distribution of local predominant wind directions (PWD) on high ozone days is useful in describing source/receptor orientations which can be predicted by a set of ozone conducive meteorological variables: daily maximum temperature, morning (7-10 a.m.) average wind speed, afternoon (1-4 p.m.) average wind speed, and midday (10 a.m. - 4 p.m.) average relative humidity. By better characterizing local mixing heights, PAMS data help better define PWDs with accompanying high ozone forming potential. The frequency distribution of the PWDs (including near calm conditions) on high ozone days may help to identify which type of ozone episodes (i.e., stagnation or transport or a mixture) is most often observed (i.e., representative) locally. Further insight into potentially significant pollutant transport may be

possible from flux analysis using PAMS upper air wind measurements (e.g, leading to detection of a nocturnal jet, such as that shown in Figure 1-1) (Lindsey, 1995). Trajectory analyses, like that shown in Figure 1-2, are useful in classifying episodes according to source/receptor orientation (Lindsey, 1995).

1.2.3   Identifying Critical Circulations Associated with High Ozone Events

Very high local ozone concentrations observed in episodes are often heavily influenced by certain meso-scale circulations embedded in large, stagnant synoptic systems. Most of these meso-scale systems are topographically induced, as in sea/lake breeze circulations and mountain-valley flows.   Under a stagnant, synoptic-scale high pressure system, these localized meso-scale systems become a major mechanism for mixing, dispersing, and transporting pollutants. For example, in coastal areas, observations often suggest that local high ozone concentrations generally coincide with the sea breeze convergent zone. Although a single PAMS upper air monitor (either a RASS or a radar profiler) may not be sufficient to resolve these meso-scale circulations, it could still provide some valuable information in the integrated analysis of the episode with surface meteorological and air quality data. An example of using the PAMS-like data in an integrated analysis of the influence of sea breeze recirculation on Houston high ozone events is illustrated in Figures 1-3 to 1-6 (Systems Applications International).

1.2.4   Identifying Boundary Layer Structures Associated with High Ozone Events

Since meteorological conditions are quite distinctive on high ozone days, it is not surprising that the structure of the atmospheric boundary layer in which the pollutants are mixing, reacting, and dispersing would be quite different from those on non-episodic days. The boundary layer profile of temperatures, winds, humidity, and mixing heights would have a direct impact on surface ozone concentrations. Due to the advanced technical capability of the instrumentation employed, the hourly winds, temperature, humidity, and reflectivity (C_n^2) data collected by the PAMS RASS and/or radar profilers add more detailed local information to the routine (and often remote) NWS rawinsonde observations and thus increase our understanding of the boundary layer structure on high ozone days. Figure 1-7 is an example showing the diurnal mixing height change derived from C_n^2 (Dye et al., 1995).

1.3   EPISODE CHARACTERIZATION USING AIR QUALITY MEASUREMENTS

PAMS data play a critical role in characterizing air quality episodes used in photochemical modeling and in the design of cost effective control measures.   Since NAMS and SLAMS were originally deployed, monitor siting technology has improved considerably such that newly located PAMS stations are meeting various objectives (e.g. maximum ozone concentration levels) with greater assurance and accuracy than previously possible. More importantly, precursor data allows

characterization of episodic events by species mix which may be used to assess the likely impact of particular strategies on ozone levels. Moreover, measured species data coupled with on-site meteorological data will lead to improved statistical models used to characterize ozone formation potential and relative frequency of occurrence for various episode types. The utility of precursor measurements to help characterize episode severity is expected to increase as the period of record for the PAMS measurements increases. This follows since the number of episodes which have precursor data and which are candidates for modeling or policy analysis will increase with passing time.

1.3.1   Indicators of Ozone Episodes

Ozone levels during the summer of 1995 in much of the US are reported to be somewhat higher than found in previous years. EPA conducted a short intensive study of the relative severity of ozone levels in 1995 compared with those in previous years using statistical methods that factored in measured meteorological conditions and urban specific ozone trends over the past decade. Based on the preliminary analysis, EPA concluded that while the summer of 1995 was unusually warm, overall conditions for elevated ozone were not atypical compared with previous years having high ozone levels. While most of the ozone measurements used in this study were from NAMS/SLAMS sites, a significant number of sites (~ 50) were operating PAMS ozone sites. One of the findings from the study was the preponderance of PAMS ozone sites (47 of 50) that reported at least one exceedance of 120 ppb during 1993-1995 suggests that PAMS sites are well placed for detecting peak ozone levels.

1.3.2   Distinguishing among Episode Types

Relative importance of transport is an important means for distinguishing among ozone episodes. Some types of ozone episodes produce a distinctive trend in diurnal patterns that may be suggestive of transport conditions. An example of such an episode for July 20-22 of 1994 in which a smooth northeasterly time progression of the ozone peak was observed over sites in New England from Lynn, MA through Jonesport ME is displayed in Figure 1-8 (NESCAUM, 1995). In a companion plot (Figure 1-9), 8-hour moving average ozone showed a clear broadening of the plume over time (NESCAUM, 1995).

Although the maximum 8-hour concentrations declined with distance, the broadening of the plume caused exposures above 70 ppb over a larger portion of the day. Since PAMS sites are located at upwind and downwind locations, this phenomenon can be more fully investigated in PAMS urban areas. For example, plots of the observed diurnal pattern for ozone, NO, NO2 and VOC species should reveal general decay in primary species coupled with an increase in secondary species if significant transport is occurring.

1.3.3   Characterizing Precursor Species During Episodes

A major strength of the PAMS program is the requirement for speciated hydrocarbon measurements to be made on a continuous (hourly or 3-hour) basis for up to 55 targeted hydrocarbon (and 3 carbonyl) compounds. Utilizing this attribute, episodes can be better characterized through analysis of the abundance of precursors experienced during periods of elevated ozone levels. An analysis of the 1994 data from the Northeast by the NESCAUM Ambient Monitoring and Assessment Committee is used to illustrate such assessments below (NESCAUM, 1995).

For the ozone episodes 07/06/94-07/08/94 and 07/20/94-07/22/94, the abundance of targeted VOCs from the five PAMS sites in the Northeast is shown in Figure 1-10 (NESCAUM, 1995). The study revealed that the most prevalent species are remarkably consistent across the region for the time period of interest. Seven compounds (i.e., isopentane, toluene, propane, ethane, n-butane, m&p-xylene and n-pentane) were found to be among the ten most abundant species at all sites except the Cape Elizabeth site. Even for this site, five of the seven species were among the ten most abundant. Interestingly, the biogenic isoprene was found at significant concentrations at four of the sites[1]. The report concluded that the pattern of similar abundances throughout the urbanized portion of the region (and to some extent, the remote portions such as Cape Elizabeth) was perhaps to be expected given the ubiquity of mobile sources.

The results above were further compared to data for five Northeastern cities from the 6:00 to 9:00 AM time period from a previous study (Wixtrom, R.N., et al.). As shown in Table1-1, six of the seven compounds of greatest abundance for the 1994 PAMS data were also the most prevalent for these five cities (NESCAUM, 1995). An extensive study of speciated VOCs in the Los Angeles area (Lurmann, F.W. , et al.) found "the same seven anthropogenic compounds to be most abundant and, with the exception of propane, in virtually the same rank of occurrence."

Volatile organic compounds react at different rates and with different reaction mechanisms due to their variations in their chemical composition and structure. As a result, VOCs differ significantly in their potential to form ozone. The use of incremental reactivities of VOCs provides a way to avoid an oversimplification of treating abundance estimates of all VOCs as equivalent. Incremental reactivity allows analysis of the effect of changing the concentration/abundance of a VOC on ozone formation. In the NESCAUM analysis, the maximum incremental reactivity (MIR) scale developed by Carter was used to show the relative ozone forming potential of the various VOC and carbonyl species (Carter, 1994). The hourly average abundances (calculated from the PAMS data) were scaled by the MIRs and displayed as

---

[1]At the East Hartford, CT site the isoprene concentrations were thought to be underestimated and compound misidentification is suspected of being reason for the lower abundance.

(potential) ppbv of ozone. The results are displayed graphically in Figure 1-11 (NESCAUM, 1995). The report noted that, although acetone and isopentane were found to be quite abundant, "their low reactivities result in rather low ozone forming potential". In addition, the increase in the relative significance of formaldehyde following the application of the MIR is highlighted. Although these data are preliminary and cover a short time span, the study concluded that the results were expected to be "a valid snapshot of upper limit conditions that typically exist in the Northeast" during periods of high ozone.

1.3.4   Assessing Airmass Aging Using PAMS Precursor Data

The same analysis of 1994 PAMS data from the Northeast by the NESCAUM Committee effectively used a comparison of ratios of VOC species to illustrate the effects of airmass aging. The series of graphics, Figures 1-12 through 1-17, from this study are used to describe the results (NESCAUM, 1995). Figure 1-12 displays the estimated benzene/toluene (B/T) and xylene/toluene (X/T) ratios based on 1990 Atlanta source profiles and source mix. Figure 1-13 shows the measured hourly ratios of benzene and m/p-xylene to toluene from the urban (type 2) PAMS site in E. Hartford, CT during the July, 1994 episode periods.

The scatter plot in figure shows that toluene levels at E. Hartford were highly correlated with both benzene and m/p-xylenes. This is consistent with a hypothesis that all three are primarily emitted by mobile sources. The B/T and X/T ratios at E. Hartford are also consistent with the predicted ratios displayed in Figure 1-13 derived from the Atlanta source profile data - suggesting that the E. Hartford source profiles and source mix are consistent with those in Atlanta (NESCAUM, 1995). Figure 1-14 shows the hypothetical effect of aging on Atlanta B/T and X/T ratios, while Figure 1-15 shows the measured B/T and X/T ratios for the rural type 3 PAMS site in Stafford, CT site (downwind of E. Hartford) during the July, 1994 episodes (NESCAUM).

The scatter plot in Figure 1-15 shows that the B/T ratio at the downwind, rural Stafford site has increased and the X/T ratio has decreased in comparison to the urban E. Hartford site shown previously in Figure 1-13 (NESCAUM). This is consistent with the predicted effect of airmass aging, as the more reactive species are differentially removed during transport. The points plotted in Figure 1-15 also exhibit greater scatter (B/T and X/T correlations are poorer) than Figure 1-13. While common (motor vehicle-related) sources are still anticipated to be a predominant cause of benzene, toluene and m/p-xylenes at the Stafford site, the species inter-correlations are diminished during transport, as the degree of aging depends on variable factors such as wind speed, wind direction, solar radiation, NO_x, etc.

Figures 1-16 and 1-17 show the relationships between toluene and m/p-xylenes at E. Hartford and Stafford, with different symbols to distinguish between daytime and nighttime samples (NESCAUM). At the urban E. Hartford site, there is relatively little difference in the X/T

ratios between nighttime samples (when reactivity is minimal) and daytime samples (when reactivity is maximal). This is consistent with a predominant, continuous influence of fresh, local, motor vehicle-related emissions at this site.

At the rural Stafford site (Figure 1-17), there's a more distinct difference between the daytime and nighttime X/T ratios. Nighttime ratios show a stronger correlation, and a slope similar to the predicted value of 0.6 for fresh emissions (and East Hartford's). For daytime samples at Stafford, there's a clear downward shift in the X/T slope (and a much poorer X/T correlation). This is consistent with a predominant influence of transported, motor vehicle-related emissions, which are photochemically aged (in a highly variable way) during the day, but which remain relatively unaged in the absence of sunlight.

1.3.5   Temporal Variation in PAMS Precursor Data

The typical daily patterns of hourly VOC concentration data and the comparison of these patterns with those of other PAMS areas, with national averages, and with historical data from this area are useful analyses to undertake to determine the unique source contributions from a particular area and changes in the composition of the urban area's ambient air quality. Average diurnal patterns or profiles are calculated by computing the average of all samples collected during each hour of the day. Most PAMS VOC species exhibit well defined diurnal cycles (or average values over time) which reflect source activity (e.g., traffic patterns), familiar daily meteorological patterns, and photochemical activity.

The first example result plotted in Figure 1-18 displays the average diurnal profiles of m,p-xylene and isoprene for six Northeastern PAMS sites for two ozone episodes in July 1994 (NESCAUM, 1995). The diurnal pattern for m/p xylene is similar to a number of other reactive, anthropogenic VOCs (toluene, o-xylene, isopentane, etc.), with emissions generally dominated by automotive-related sources. Isoprene is emitted predominantly by deciduous vegetation, as a function of solar radiation and temperature. As a biogenic compound, isoprene has a unique diurnal pattern which is distinct from that of the anthropogenic VOC species. While the morning (6:00 to 9:00 AM) levels and reactivities of m/p xylene and isoprene are quite similar, the reactive anthropogenic pollutants are generally depleted rapidly during the day. Although isoprene also reacts rapidly, its rate of production exceeds its rate of destruction during mid-day.

The next example result, Figure 1-19, introduces a variation on the basic diurnal pattern technique depicted above by using box plots for each hour rather than displaying simple means. Note that each panel contains 24 box plots corresponding to each hour of the day. The figure contains diurnal patterns for 1993 PAMS data on acetylene, olefins, toluene, ethylene, xylene and isoprene from Baltimore's Site #2 (Cox, 1995).   The organic species for this analysis clearly indicate the typical diurnal trends for anthropogenic VOCs described above. Median values for all

species, except isoprene, show a tendency for higher morning and evening concentrations. These patterns are interpreted as follows:

- the morning maximum is associated with high emissions and limited mixing;
- the mid-day minimum is associated with decreased mobile source emissions, increased mixing due to rapid growth of the daytime boundary layer, and increased reaction rates due to higher temperatures and maximum solar radiation;
- the early evening maximum is associated with gradual build up of emissions in the boundary layer over the course of the day, increased mobile source emissions during the afternoon culminating in an early evening commute traffic peak, and decreased mixing.

The diurnal profile for isoprene reflects the fact that biogenic emissions are a strong function of temperature and solar radiation and are short-lived in the atmosphere.

The final example results for this section on diurnal patterns are from the analysis of 1993 PAMS data from Houston, Texas (Stoeckenius, November 1994). Figures 1-20a through 1-20c present separate calculations of the diurnal profiles of species concentrations for weekdays and weekends; an effort to identify the impact of differences in emissions between these two groups of days[2]. Comparisons of profiles between species and between weekdays and weekends can reveal much about underlying meteorological, chemical, and emission factors. The vertical bars in these figures indicate the 95 percent confidence limits for the mean value in each hour. Thus, pairs of hours with non-overlapping bars have significantly different mean concentrations.

At the Galleria site, weekday concentration profiles for all species except isoprene exhibit a strong morning peak. Morning peaks of acetylene, ethylene, toluene, olefins, and xylenes are noticeably lower on weekend mornings; these differences are consistent with reduced morning mobile source emissions on weekends. The lowest concentrations occur between 9:00 AM and 1:00 PM for all species (except isoprene) and the levels rise again to a second peak by early evening. These patterns are similar to those found in other analyses by other investigators.

---

[2]Monday and Friday holidays were grouped in with weekends. Holidays falling in the middle of the week were left out of this analysis since traffic patterns and business activities on these days tend to differ from regular weekdays or weekends. Mean values were not computed for hours for which data was not available for at least two-thirds of sampling days. This restriction primarily affected the hours around midnight (hours 11, 0, and 1) which were sampled less frequently than other hours. A cubic spline fit was used to generate the smooth curves in these figures; the dots along the curves represent the actual hourly averages and are plotted at a position corresponding to the mid-point of the hour they correspond to. This plotting technique accentuates the diurnal patterns making graphs that are easy to read but one must be careful not to over interpret the peaks and troughs that may be shown as occurring between the actual hourly average values.

Diurnal profiles (based on weight percents) for the same VOC species above are presented in Figures 1-21a through 1-21c (Stoeckenius, November 1994). For the most part, the profiles are similar to the concentration profiles in Figures 1-20a through 1-20c. Some differences are evident, however. Toluene, olefins, and xylene weight percents exhibit smaller diurnal variations than their concentration counterparts although the pattern is basically the same.

1.3.6   Statistical Models of Relationship Between Ozone and Precursors

In the context of the PAMS program, the use of regression analysis involves developing empirical models to statistically describe the relationship (or potential relationship) between independent parameters (ozone precursors and meteorological data) and the dependent parameter (ozone). This analytic technique takes advantage of one of the unique characteristics of the PAMS data: the concurrent measurement of VOCs, meteorological parameters, nitrogen species and ozone. The results of these analyses (explanation in the variability of daily ozone maxima based on meteorological and precursor data) can lead to invaluable insights about the design of effective control strategies for the area analyzed.

However, this analytic technique and the example results which follow should only be considered as "exploratory" given the early stages of the PAMS program and the remaining further investigations that are required to verify and understand these initial evaluations. The reader is cautioned not to make inferences as to the nature of cause-effect relationships between the VOC species included in the regression models and downwind ozone (i.e., it is premature to predict or conclude how ozone levels would behave if concentrations of the included VOCs were reduced) based on these analytic results. These analyses should be viewed as another means to begin using PAMS data to develop a more complete understanding of the ozone phenomenon in PAMS cities. Finally, it is important to remember that these models have been constructed (to date) without selecting variables for inclusion in the model based on their hypothesized physical relationship to ozone.

The examples presented below demonstrate the technique using the 1994 PAMS data from the Philadelphia, PA[3]. The daily maximum ozone concentration, the dependent variable, was calculated from the downwind ozone data (Site#3). Data for the independent variables (i.e., VOCs, nitrogen species, meteorological parameters) were taken from the upwind urban site (Site #2). Several indicators for the independent variables were calculated (averages and maxima for selected intervals: three, six, twelve and twenty-four hour periods during the day of interest or lagged for earlier days). The relationships between all of these indicators and the dependent variable were then analyzed and those with the strongest correlations were then selected for

---

[3] The ozone precursor and meteorological data included in the model are from the PAMS Site #2, and the downwind ozone data are from the PAMS Site #3.

inclusion in the regression model. Figure 1-22 uses a scatter plot to display how well the initial model, Model 1, estimated the observed daily maximum ozone values. This model utilized only meteorological parameters: temperature (mean for interval 12:00 PM to midnight) and wind speed (maxima for interval 6:00 to 9:00 AM).   The R-square value for this model is 0.45 reflecting that the model using only meteorological parameters explains approximately 45% of the variability of downwind ozone levels.

The results presented in Figure 1-23 reveal an improvement in the model's "goodness of fit" (R-square) with the inclusion of VOCs. The following hydrocarbon species have been included with the earlier meteorological data in Model 2[4]:

VOC   Statistic/Interval

Isopentane   Mean/8 hr lag
3-Methyl pentane   Mean/9-12
n-Hexane   Mean/8 hr lag
n-Octane   Mean/19-22
Nitric Acid   Max/13-15
2,2-Dimethyl butane   Mean/6-9
Nitrogen Dioxide   Max/1-12
2-Methyl heptane   Mean/6-9
Ethyl benzene   Max/8 hr lag
Propane   Mean/8 hr lag

The resulting R-square value (0.84) is an improvement over that of the meteorology-only model. This example result suggests that the measurement of speciated hydrocarbons is essential for accurately representing and understanding the character of maximum ozone levels.

Neural networks provide a more flexible alternative to standard regression methods for relating dependent variables to a set of independent variables. Because neural networks are more general, they can accommodate both non-linearities and interactions among independent variables without explicit parameterizations required in non-linear regression models. While neural networks offer the potential for better prediction of the response variable, results are often difficult to interpret, mainly due to model complexity and multicollinearity of the process. Weights (regression coefficients) are typically unstable, usually "inflated" and vary considerably from true optima if local optima are found in the fitting process. Another major drawback is that extrapolation is risky: producing significant errors under certain conditions.

Crowe and coworkers have applied neural networks using data collected in the southeast Texas region at the Galleria Site near Houston. The data consisted of hourly meteorological data (net radiation, temperature, wind direction and speed, wind variation), nitrogen oxides (NO_2 and NO_x) and seven hydrocarbon species based on carbon bond 4 chemistry. Hourly ozone data were taken from a downwind site located at the Clinton site near east Houston. Three neural network

---

[4] An additional model (not presented in this report) for the 1994 Philadelphia PAMS data was constructed and is available upon request. The model incorporated a number of additional VOC species to those already utilized in Models 1 and 2.

models were developed using proprietary software (Process Insight) developed by Pavilion Technology Inc. The predictor variables for the first model (MET) consisted of the five meteorological parameters for the same hour as the ozone measurement and also for 6 time delays to account for possible effects of transport and chemical reactions.   The second model ($METNO_X$) included the same meteorological variables but added the species NO and NO_X. The third model ($METNO_XCB4$) dropped the time lagged variables (apparently unimportant) but added seven hydrocarbon species based the carbon bond 4 chemistry.

The models showed progressively better predictive capability (using the entire data set) as evidenced by increasing R^2 values from 0.70 (MET), 0.80 ($METNO_X$) and 0.91 for $METNO_XCB4$). The authors reported that selected hydrocarbon species are more sensitive predictors of hourly ozone--increasing olefins associated with decreasing ozone and increasing paraffins associated with sharply increasing ozone levels. The authors view this work as very preliminary and have neither attempted to associate any cause and effect to such relationships nor attributed any physical and/or chemical significance to their findings to date.

Capone applied neural network technology to predict down wind hourly ozone data in the Baton Rouge area using a more complicated network in which data from two downwind sites were used as predictors. The model consisted of hourly meteorological and NO_X measurements (NO2, NO and NO_x) at each site and was successful at predicting hourly ozone patterns as evidenced by graphics comparisons of the diurnal pattern and scatter plots between measure and predicted values. Interestingly Capone's model did not involve any hydrocarbon species, apparently due to a lack of relatively complete measurements in this area.

Meteorological and PAMS data for 1994 were used in an exploratory application of neural networks to predict daily maximum 1-hour ozone levels in the Philadelphia CMSA. The PAMS data were taken from an upwind Delaware site (10-003-1007) where hourly measurements are available from June-August. PAMS data included in the analysis were average TNMOC (6-9 am), olefins (6-9 am), NO (6-9 am) and NO_2 (6-9 am) and mid-day maximum hour isoprene. Daily composite meteorological variables were computed using available data from the nearest National Weather Service station. These variables included maximum surface temperature, morning and afternoon average wind speed, and mid-day average relative humidity and cloud cover.   Out of 92 candidate days, only 51 days had sufficient data to be included in this analysis.

Using commercially available software, a neural network was used to relate maximum daily ozone to the combination of five meteorological and five PAMS precursor variables. For comparison purposes, neural networks of size 1 (logistic regression--11 weights) through size 3 (a total of 37 "weights") were fit using all 10 predictor variables. Figure 1-24 is a scatter plot of the log of the observed vs predicted ozone using for the smallest sized network. The fit is relatively good ($R^2 = 0.68$) which is not surprising since the five meteorological variables alone are known to

be good predictors of ozone forming potential in many U.S. urban areas. The R^2 using just the five meteorologically variables is nearly as good (0.64) suggesting that the precursor species provides only a marginal improvement in the fit.

The coefficients (weights) of the independent variables from this single node neural network may be interpreted in a similar manner as with ordinary regression models. For example, the coefficients for temperature are positive suggesting that increases in temperature are associated with higher ozone. Conversely, the coefficients for wind speed and cloud cover are negative suggesting that increases in winds and cloud cover are associated with lower ozone levels. Of the species variables, olefins and mid-day isoprene were negatively related to ozone although the magnitude of the coefficients were small relative to their standard errors.

Figure 1-25 show a similar plot of fitted ozone from a neural network of size=3. Although the R^2 statistic has increased considerably (0.89) the number of fitting parameters is large relative to the number of observations. Also, the relative complexity of the model makes interpretation of the coefficients difficult. For example, coefficients for temperature are both positive and negative in the three linear inputs feeding the hidden layer. Although the net effect of temperature is positive, it is difficult to interpret the relative role that temperature (and other variables) have within each layer. As more PAMS data becomes available, the stability of larger sized neural networks can be better assessed.

Clearly, the process of model building is very much an art at this point and will require close coordination among dispersion modelers, atmospheric scientists, and data analysts. Neural networks may help establish a practical upper bound on the predictive ability of statistical models and provide insight into reasonable model structures that may be more interpretable. Hopefully, the process will lead to development of physically meaningful input parameters that will help simplify the structure of these models and lend credibility to their potential applications. For example, better predictive models should provide better selection of episodes for modeling and more accurate assessment of the severity of episodes used in model based attainment demonstrations. Also, properly structured empirical models may provide supportive information regarding the potential effectiveness of emission control strategies including the relative benefits of VOC/NO_x emission reductions on ozone levels.

Crowe and coworkers applied neural network methods to predict hourly ozone levels using PAMS like data taken at Texas air monitoring sites during 1990-1994. They report R^2 statistics on the order of 0.75 when meteorological data alone are used in the model. When hourly precursor concentrations of selected VOC's, NO_x are included, R^2 statistics increase to approximately 0.85 suggesting that precursor species are important predictors of ozone levels.

1.4   REFERENCES

Bruntz, S.M.; Cleveland, W.S.; Graedel, T.E.; Kleiner, B.; and Warner, J.L. "Ozone Concentrations in New Jersey and New York: Statistical Association with Related Variables." Science 186, 1974.

Capone, R.L. Presentation to OAQPS, "Predicting Downwind Air Quality with a Neural Network", March, 1996.

Carter, W.P.L. "Development of Ozone Reactivity Scales for Volatile Organic Compounds." Journal of Air and Waste Management Association 44, 1994: 881-898.

Chu, S. H. "Coupling High Pressure Systems and Outbreaks of High Surface Ozone Concentration."   Proceedings of the 80th APCA Annual Meeting. 1987: 87-113.5.

Chu, S. H. "Meteorological Considerations in Siting Photochemical Pollutant Monitors." Atmospheric Environment 29, 1995: 2905-2913.

Chu; S. H. and Doll, D. C. " Summer Blocking Highs and Regional Ozone Episodes."   Preprints of the Seventh Joint Conference on Applications of Air Pollution Meteorology with AWMA, American Meteorological Society 1991: 274-277.

Cox, W.M. "A Workbook for Exploratory Analysis of PAMS Data." June 1995.

Cox, W. M. and Chu, S.H. "Meteorologically Adjusted Ozone Trends in Urban Areas: A Probabilistic Approach." Atmospheric Environment 27B, 1993; 425-434.

Cox, W. M. and Chu, S.H. "Assessment of Interannual Ozone Variation in Urban Areas from a Climatological Perspective." Atmospheric Environment 30, 1996: 2615-2625.

Crowe, W. and DeFries, T.H. Presentation to OAQPS, "Use of Observation Based Models to Predict Ambient Ozone Levels." February 27, 1996.

DeFries, Timothy H. Neural Network Modeling of Ambient Ozone Using the South East Texas Regional Planning Commission Ambient Air Monitoring Data Set. Radian Corporation, June 12, 1995.

Dye, T.S.; Lindsey, C.G.; and Anderson, J.A., "Estimates of Mixing Depths from 'Boundary Layer' Radar Profilers." Preprints of the 9th Symposium on Meteorological Observations and Instrumentation Charlotte, NC, March 27-31, 1995.

Hosler, C.R. "Low-level Inversion Frequency in the Contiguous United States." <u>Mon. Wea. Rev.</u> 89, 1961: 319 - 339.

Lamb, B.; Guenther, A.; Gay, D.; and Westberg, H. "A National Inventory of Biogenic Hydrocarbon Emissions." <u>Atmospheric Environment</u> 21, 1987: 1695 - 1705.

Lindsey, C.G.; Dye, T.S.; Roberts, P.T.; Anderson, J.A.; and Ray, S.E. <u>Meteorological Aspects of Ozone Episodes in Southeast Texas</u>. Paper No. 95WP96.02 presented at the 88th Air & Waste Management Association Annual Meeting, San Antonio, TX, June 18-23, 1995.

Lindsey, C.G.; Dye, T.S.; Blumenthal, D.L.; Ray, S.E.; and Arthur, M. <u>Meteorological Aspects of Summertime Ozone Episodes in the Northeast</u>. Paper FA 5.8 presented at the 9th Joint Conference on the Applications of Air Pollution Meteorology with AWMA at the 76th American Meteorological Society Annual Meeting, Atlanta, GA, January 28-February 2, 1996.

Ludwig, F.L.; Reiter, E.; Shelar, E.; and Johnson, W.B. <u>The Relation of Oxidant Levels to Precursor Emissions and Meteorological Features, Part 1: Analysis and Findings, Final Report</u>. EPA Contract 68-02-2084. SRI International, Menlo Park, CA. EPA Report No. 450/3-77-022a, 1977: 153.

Lurmann, F.W. and Main, H.H. <u>Analysis of Ambient VOC Data Collected in the Southern California Air Quality Study. Report to the California Air Resources Board</u>. Sonoma Technologies, Inc., Santa Rosa, CA. 1992.

Northeast States for Coordinated Air Use Management (NESCAUM), The Ambient Monitoring and Assessment Committee. <u>Preview of 1994 Ozone Precursor Concentrations in the Northeastern U.S.</u> August 1995.

Pagnotti,V. "Seasonal Ozone Levels and Control by Seasonal Meteorology." <u>Journal of the Air & Waste Management Association</u> 40, 1990.

Robinson, E. "Some Air Pollution Aspects of the Los Angeles Temperature Inversion." <u>Bulletin of the American Meteorological Society</u> 33, 1952: 247 - 250.

Stoeckenius, T.E.; Ligocki, M.P.; Shepard, S.B.; and Iwamiya, R.K. <u>Analysis of PAMS Data: Application to Summer 1993 Houston and Baton Rouge Data, Draft Report</u>. U.S. EPA Contract 68D30019, Systems Applications International, SYSAPP-94/115d. November, 1994.
Stoeckenius, T.E.; Ligocki, M.P.; Cohen, J.P.; Rosenbaum, A.S.; and Douglas, S.G. <u>Recommendations for Analysis of PAMS Data, Final Report</u>. U.S. EPA Contract 68D30019, Systems Application International, SYSAPP94-94/011r1. February, 1994.

Systems Application International, Sonoma Technology, Inc., Earth Tech, and Alpine Geophysics. Gulf of Mexico Air Quality Study, Volume 1: Summary of Data Analysis and Modeling. Draft final report prepared for U.S. Department of the Interior, Minerals Management Service, Gulf of Mexico OCS Region, New Orleans, LA. OCS Study, MMS 94-0046, SYSAPP-95/013d, 1995.

Wixtrom, R.N. and Brown, S.L. In: Journal of Exposure Analysis and Environmental Epidemiology 2, 1992: 51. Edited by Edo Pellizzari (averages recalculated here for five Northeast cities only).

Figure 1-1.
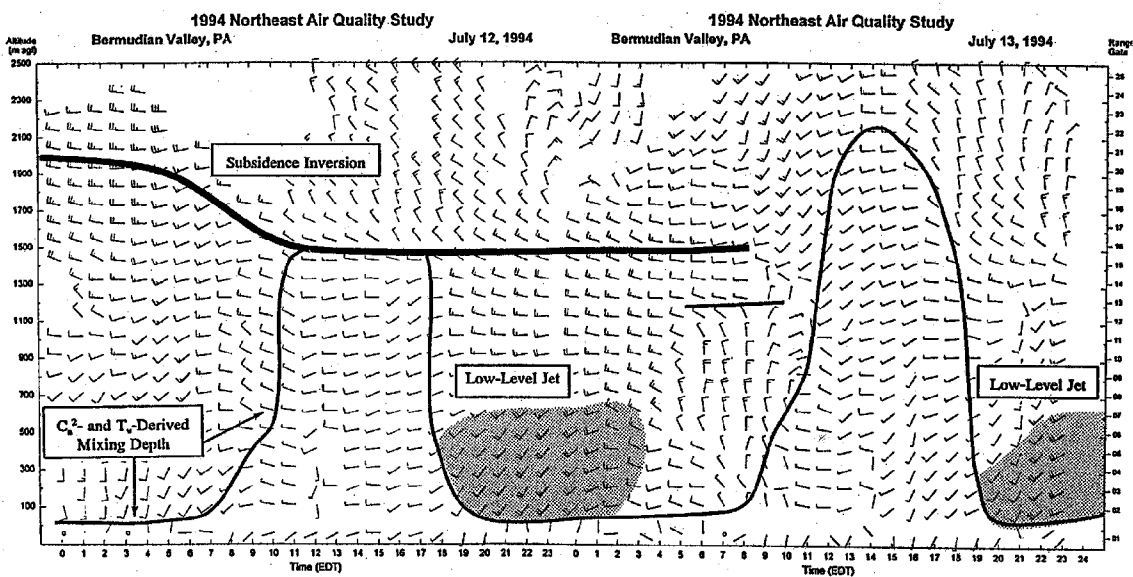
Time series cross section of winds, mixing depth, and inversion conditions measured by the radar profiler on July 12-13, 1994 at Bermudian Valley, PA. The thin solid line denotes the height of the mixed layer estimated using C_n^2 and RASS temperature data. The thick line denotes the subsidence inversion. The shaded area indicates the region of the nocturnal low-level wind maxima (Lindsey et al., 1995b).
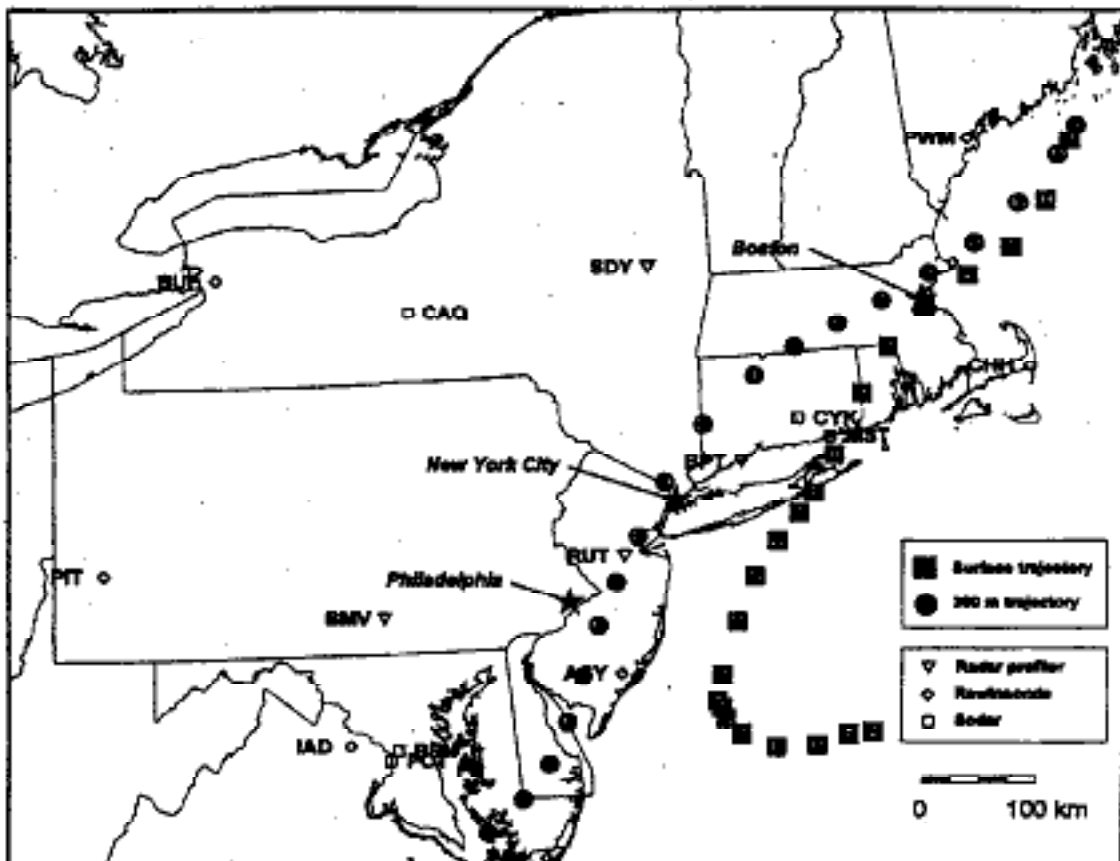
Figure 2-2. Back trajectories at the surface and 300m agl calculated from a ferry equipped with an ozone monitor (operated by the state of Maine) that recorded exceedances of the ozone standard on July 21, 1994.

Figure 3-3.
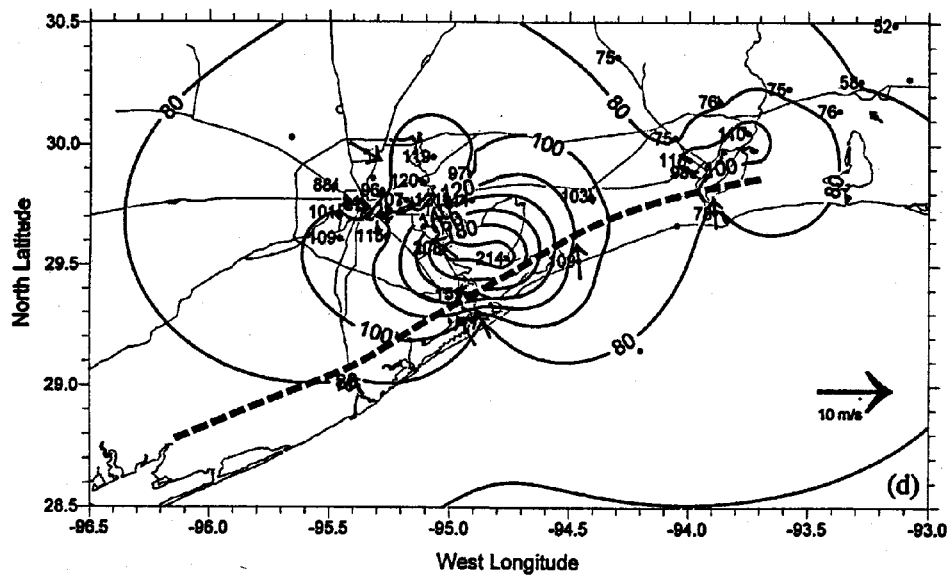
Plot of surface winds and ozone concentrations in the southeast Texas region at 1400 CST on September 8, 1993 (SAI 1995). The dashed line indicates the location of the sea breeze front.

Figure 5-4.
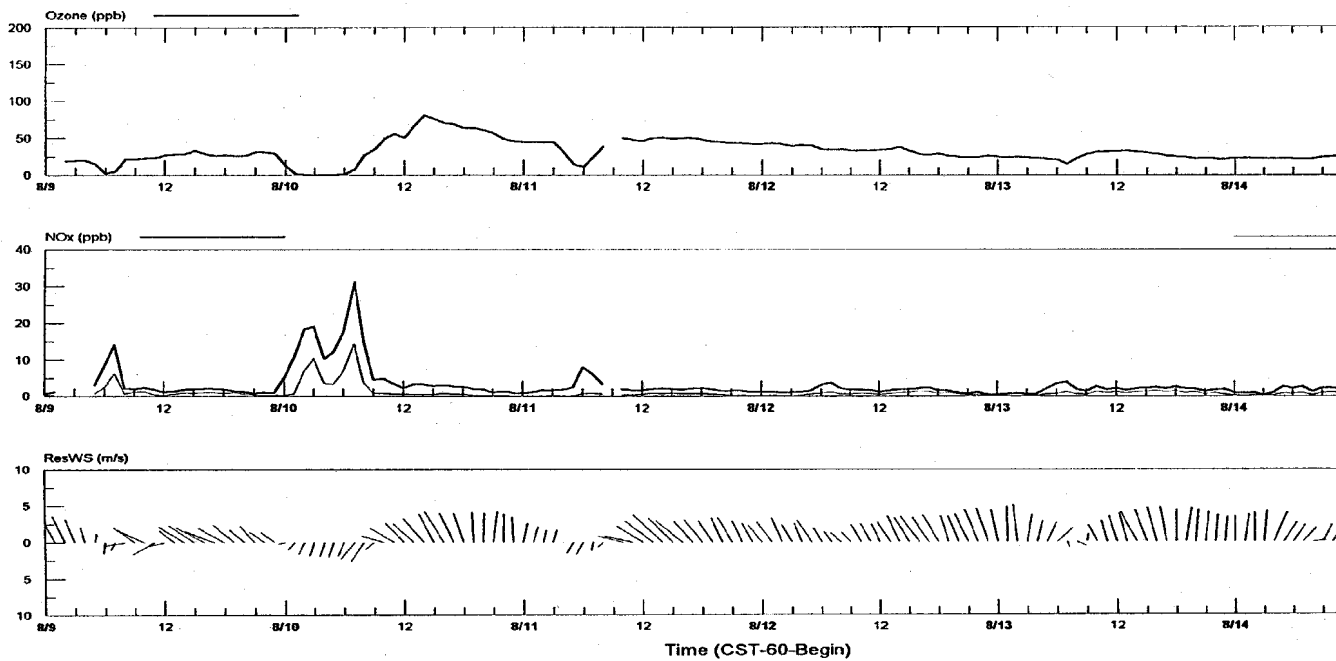
Time series of ozone, NO, NO$_x$, wind speed, and wind direction measured at the Gilcrest, TX surface monitoring period August 9-14, 1993 (SAI, 1995). When flow was offshore, high NO$_x$ and titrated ozone was observed at When the flow reversed, high ozone concentrations were observed. However, with continued onshore flow, ozo concentrations decreased.

Figure 6-5.
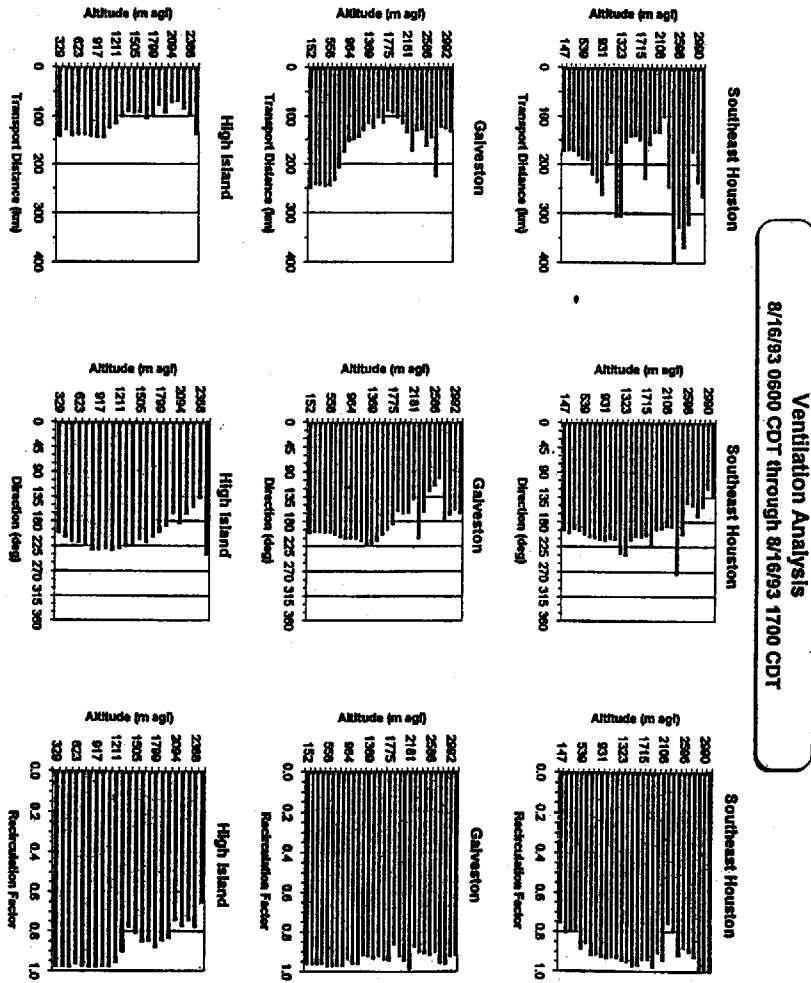
Vector-integrated transport distances, resultant wind directions, and recirculation factors (R), calculated from data collected by the southeast Houston (SEH), Galveston (GAL) and High Island Platform (HIP) radar profilers for the period 0600-1700 CDT on August 16, 1993.
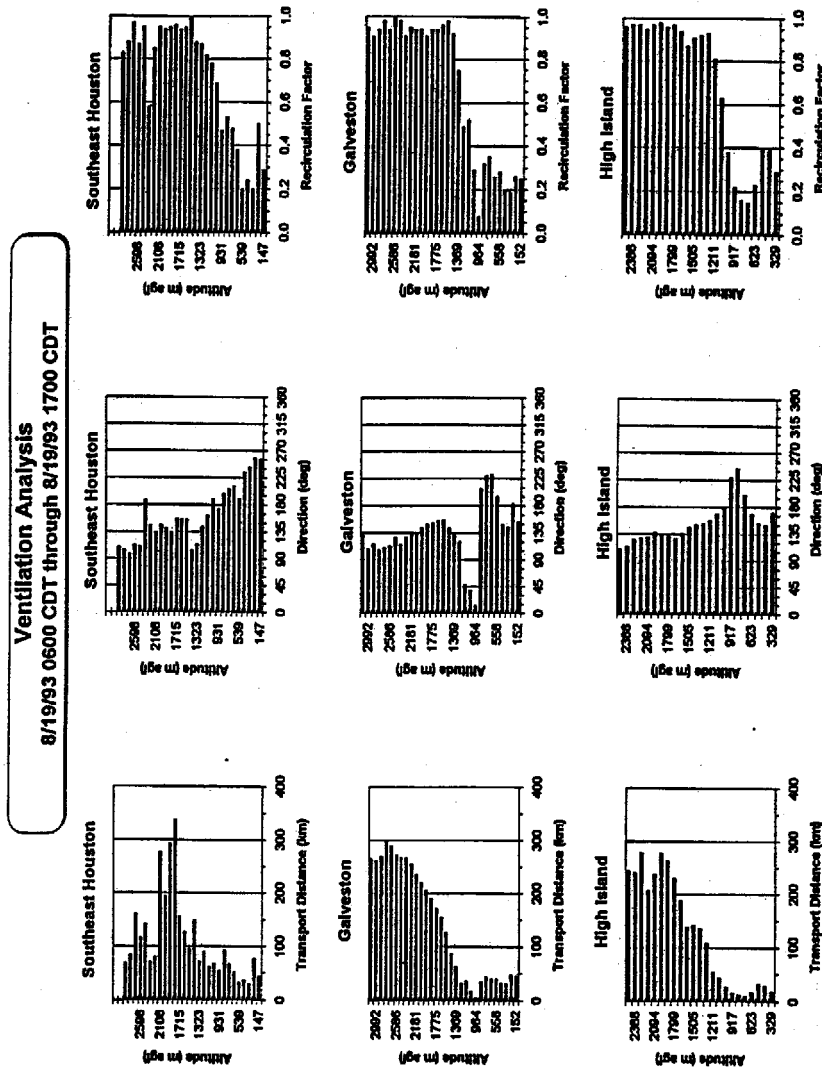
Figure 7-6. Vector-integrated transport distances, resultant wind directions, and recirculation factors (R), calculated from data collected by the southeast Houston (SEH), Galveston (GAL) and High Island Platform (HIP) radar profilers for the period 0600-1700 CDT on August 19, 1993.
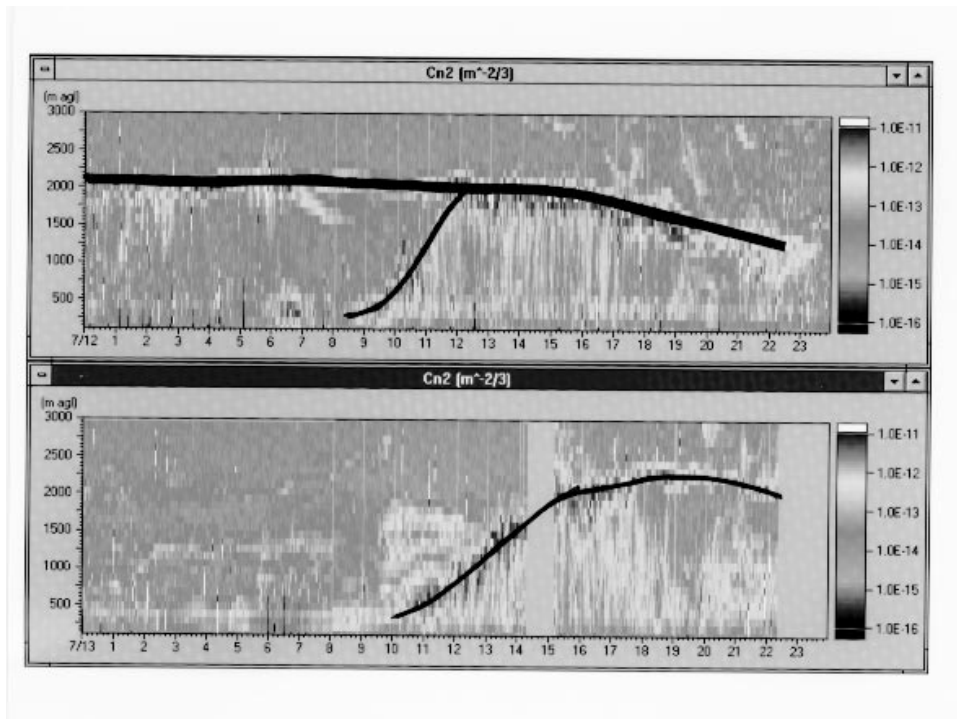
Figure 1-7.   Time-height cross-section of C_N^2 for July 12-13, 1994 at New Brunswick, NY. Thick line denotes the subsidence inversion; thin line during the day denotes the top of the mixed layer. On July 12, a subsidence inversion is shown in the profiler data as a region of high reflectivity between 1750 and 2000 m agl. The inversion limited afternoon growth of the CBL to below 2000 m. A slower growth occurred on July 13 resulting in reduced vertical mixing of precursors, and an attendant elevated afternoon ozone level.

Jonesport (86 ppb at 21:00)

Bar Harbor, ME (102 ppb at 19:00)

Isle Au Haut, ME (116 ppb at 17:00)

Port Clyde, ME (124 ppb at 17:00)

Phippsburg, ME (148 ppb at 15:00)

Cape Elizabeth, ME (148 ppb at 15:00)

Kennebunkport, ME (141 ppb at 14:00)

Rye, NH (135 ppb at 13:00)

Lynn, MA (105 ppb at 12:00)
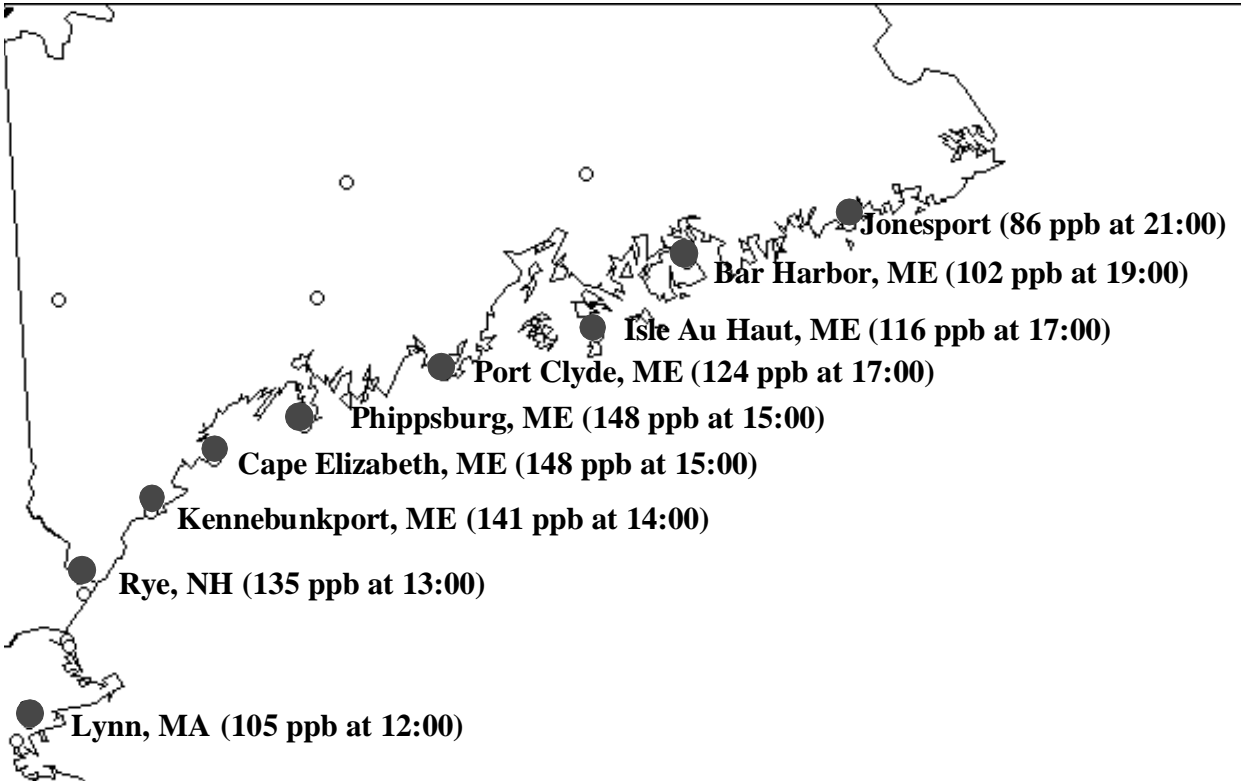
Figure 9-8. Maximum Ozone Concentrations and Hours of Occurrence, 07/21/94
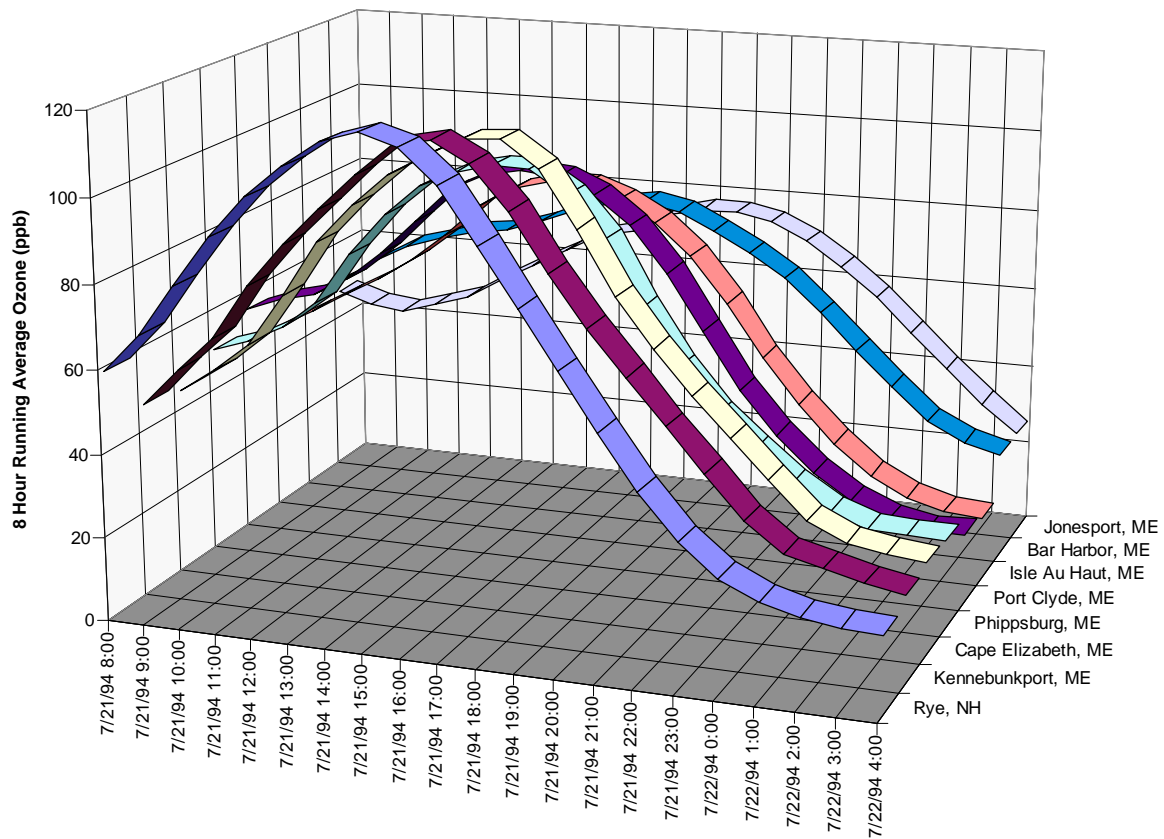
Figure 1-9.   8-Hour Moving Average Ozone Concentrations on 7/21/94
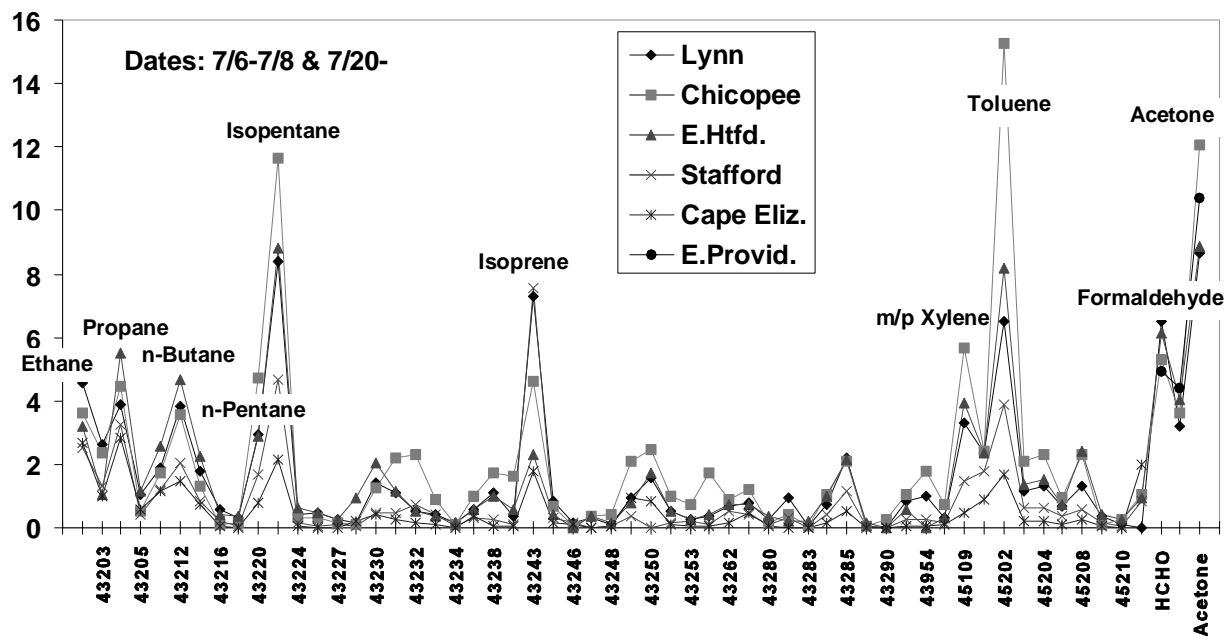
Figure 11-10. VOC Abundances for Six Northeastern PAMS Sites, July 1994.
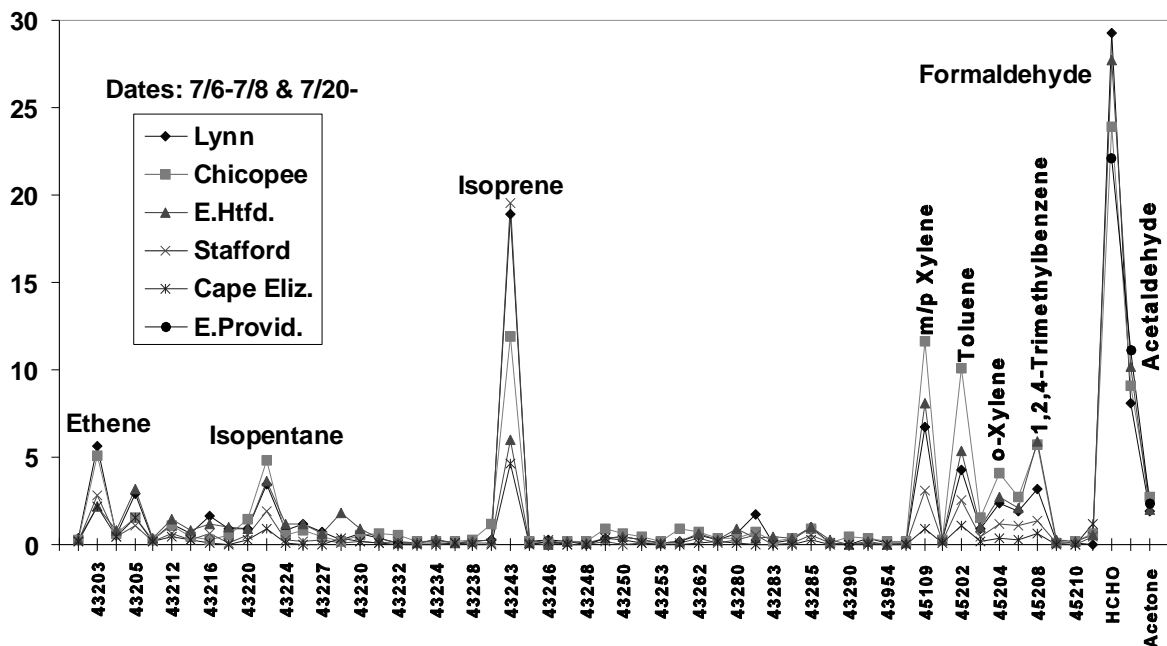
Figure 12-11. VOC Abundances Adjusted for Reactivity (Using Carter's MIRs) for Six Northeastern PAMS Sites, July 1994.
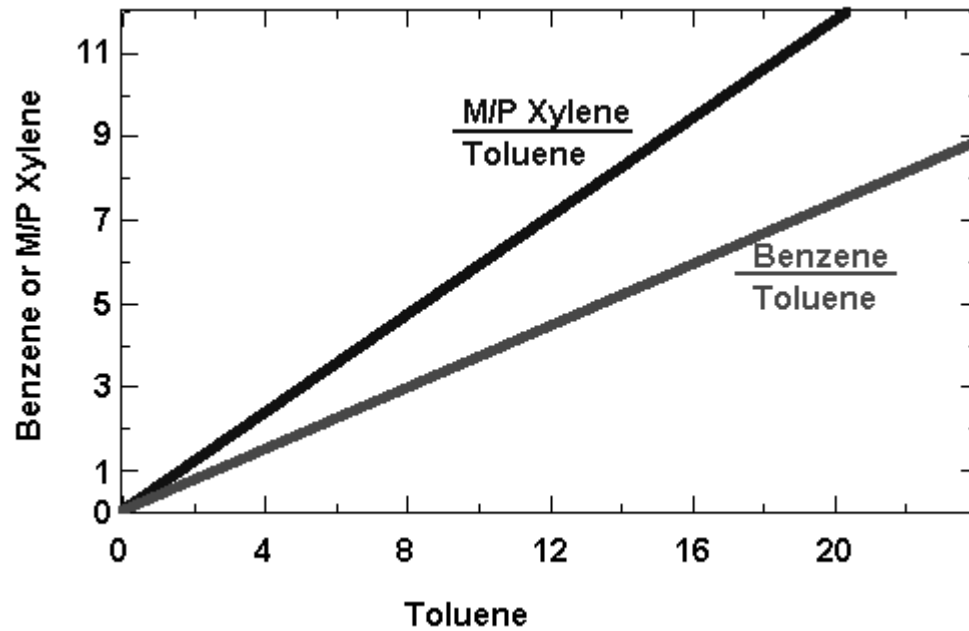
Figure 13-12. Estimated Urban B/T and X/T Ratios from Atlanta, GA Source Profiles (from Henry et al., 1994)
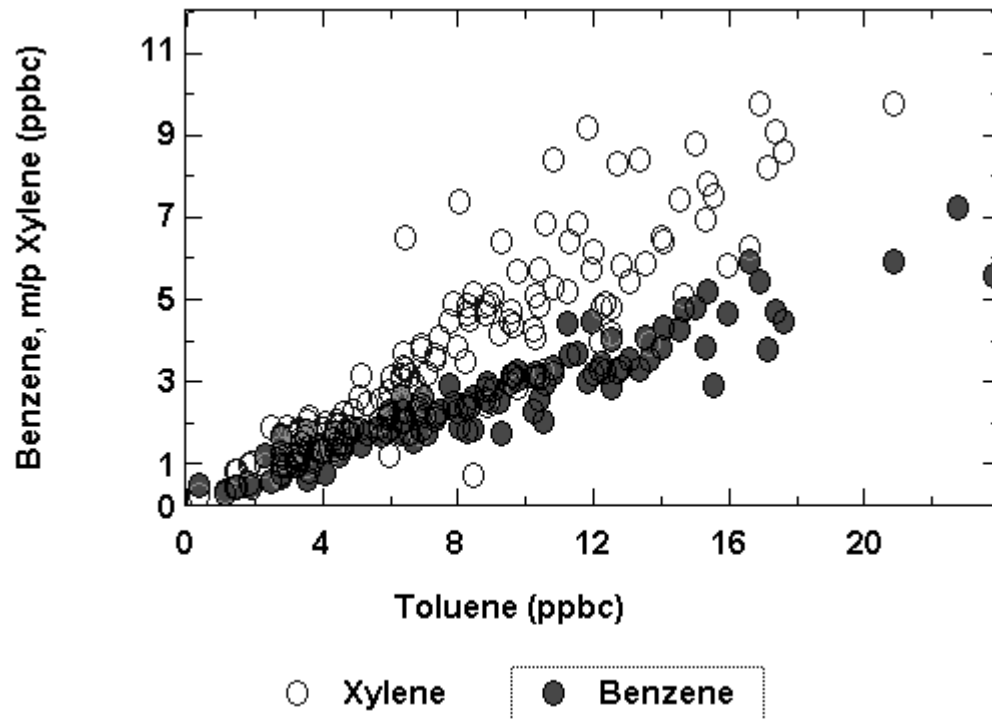
Figure 14-13. Measured Urban B/T and X/T Ratios from E. Hartford, CT PAMS Site during July 1994 Episodes
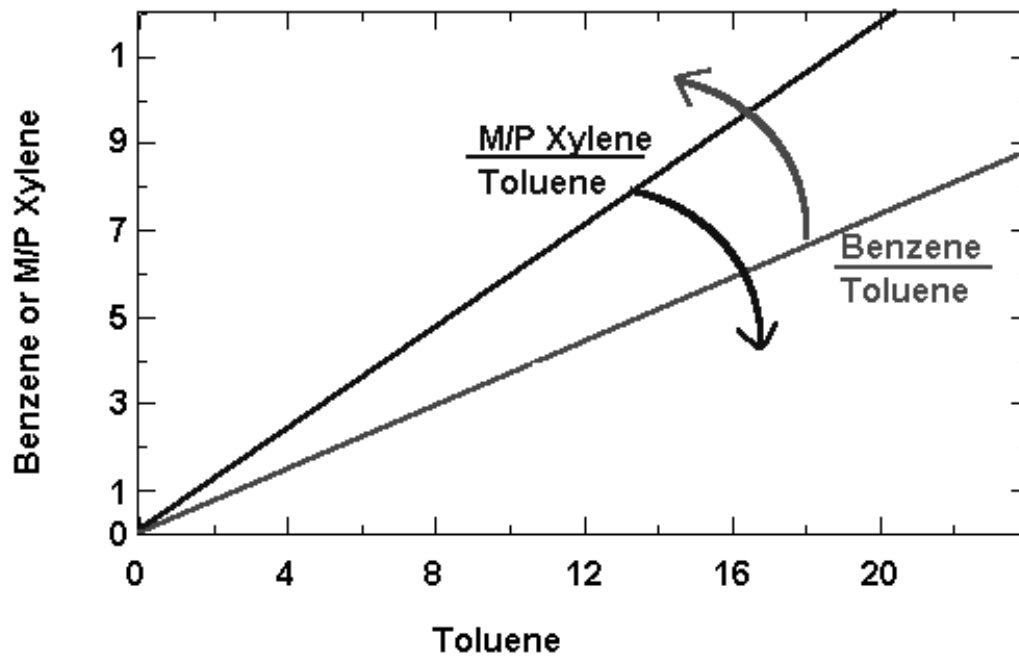
Figure 15-14. Predicted Changes in B/T and X/T Ratios at Rural Sites (Resulting from Airmass Aging)
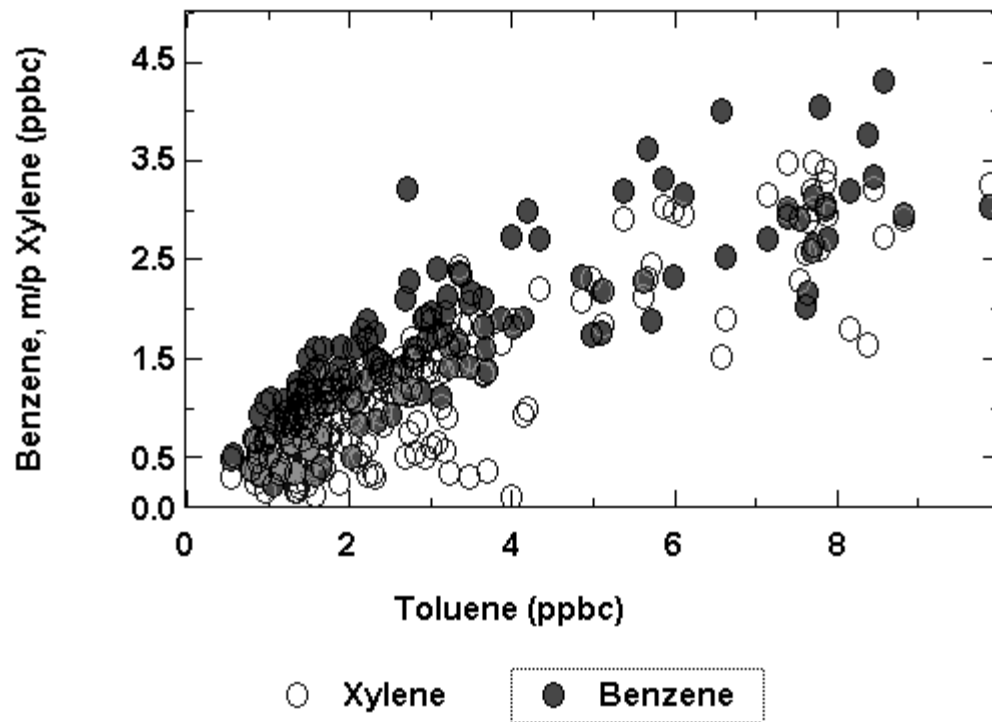
Figure 16-15. Measured Rural B/T and X/T Ratios from Stafford, CT PAMS Site during July 1994 Episodes
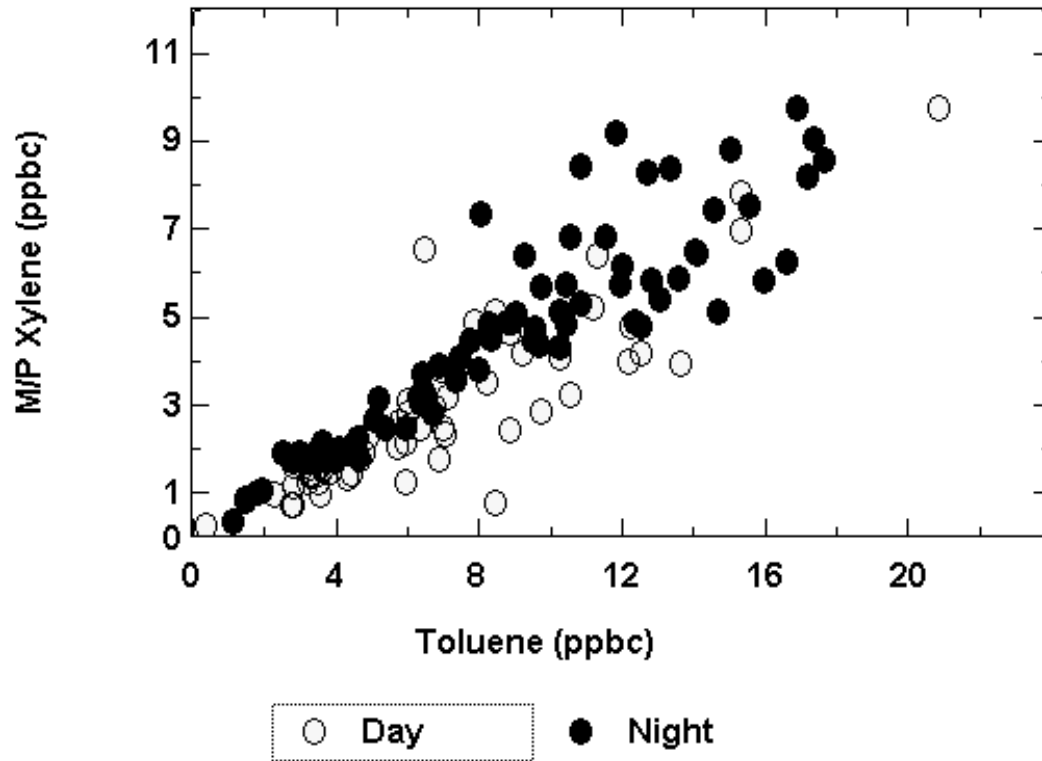
Figure 17-16. m/p-Xylene versus Toluene at E. Hartford, CT PAMS Site during July 1994 Episodes
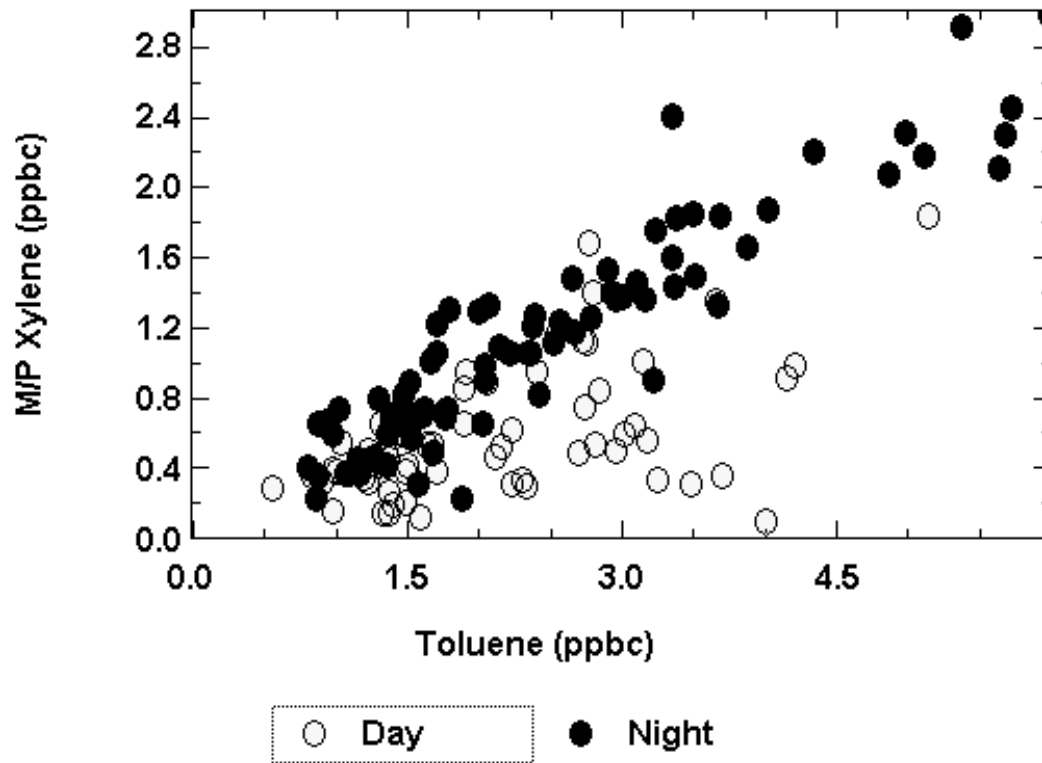
Figure 18-17. m/p-Xylene versus Toluene at Stafford, CT PAMS Site during July 1994 Episodes
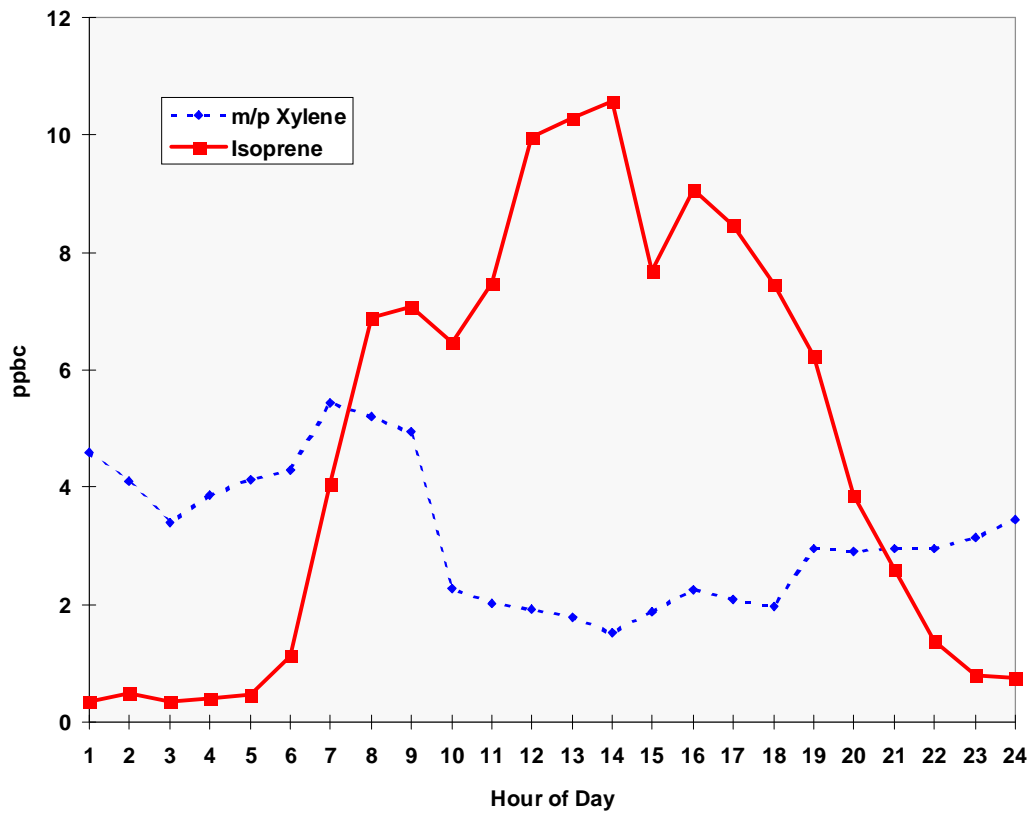
Figure 1-18   Diurnal Isoprene and m/p Xylene (averaged for 6 NESCAUM PAMS Sites for July, 1994 Episodes)
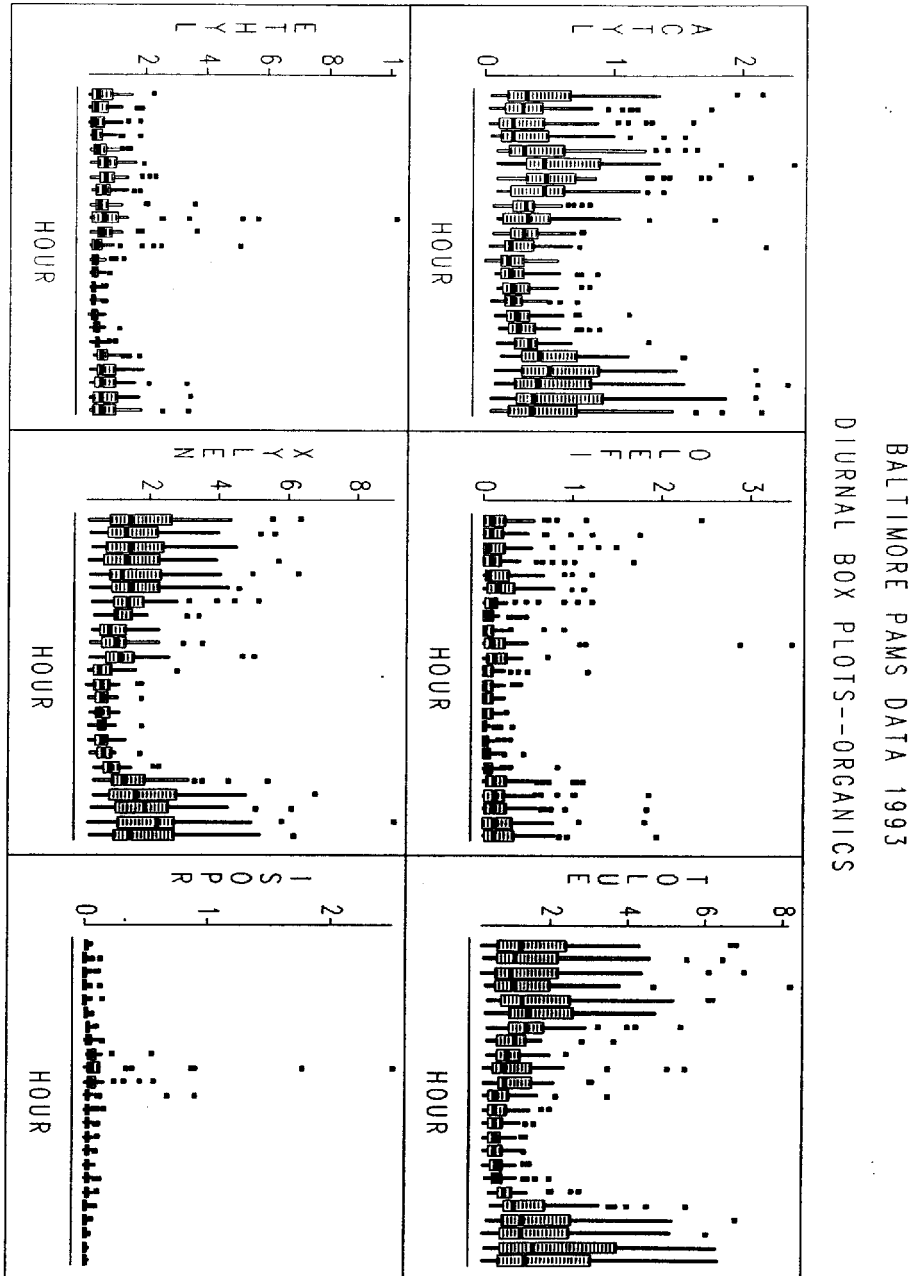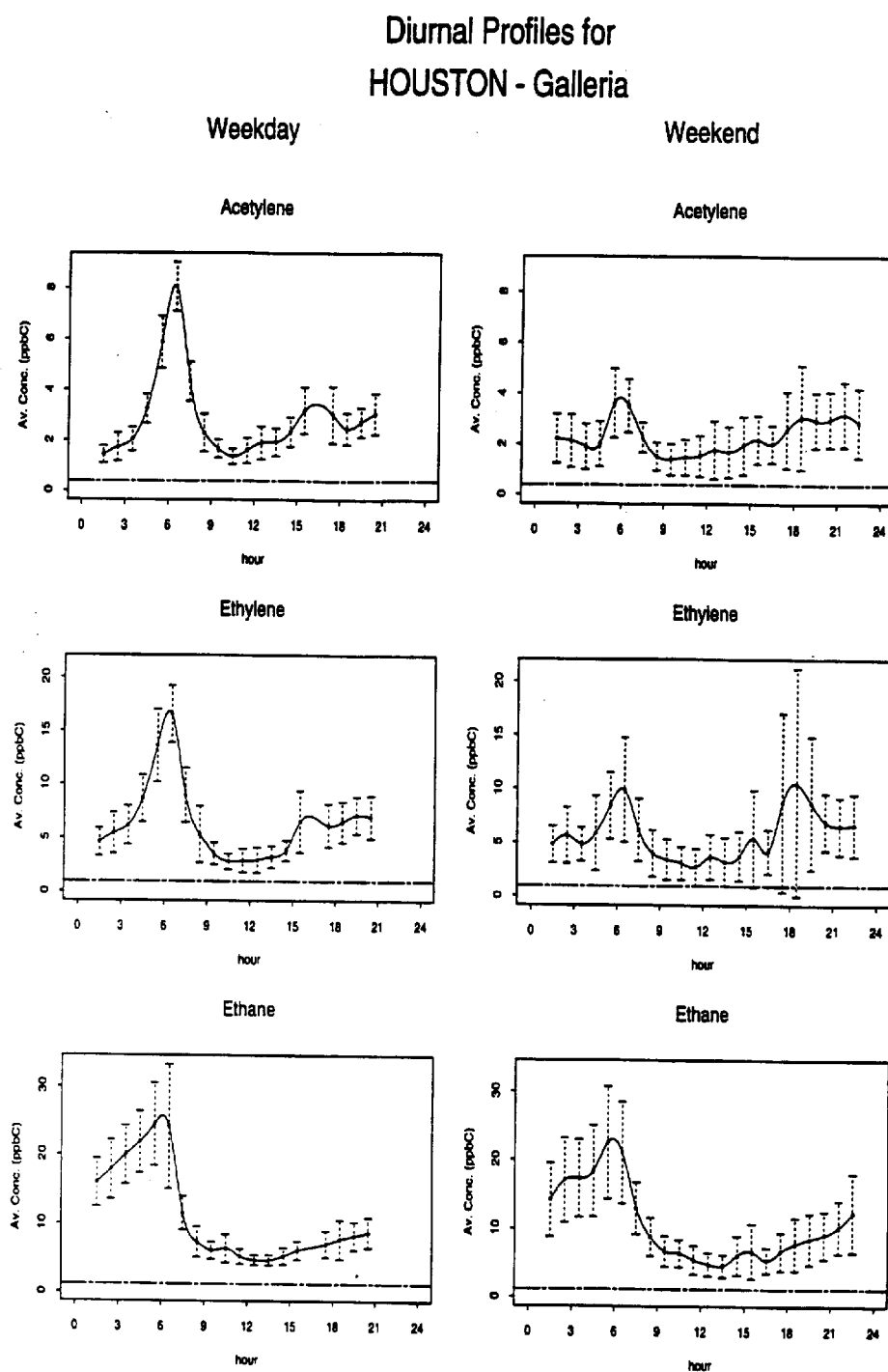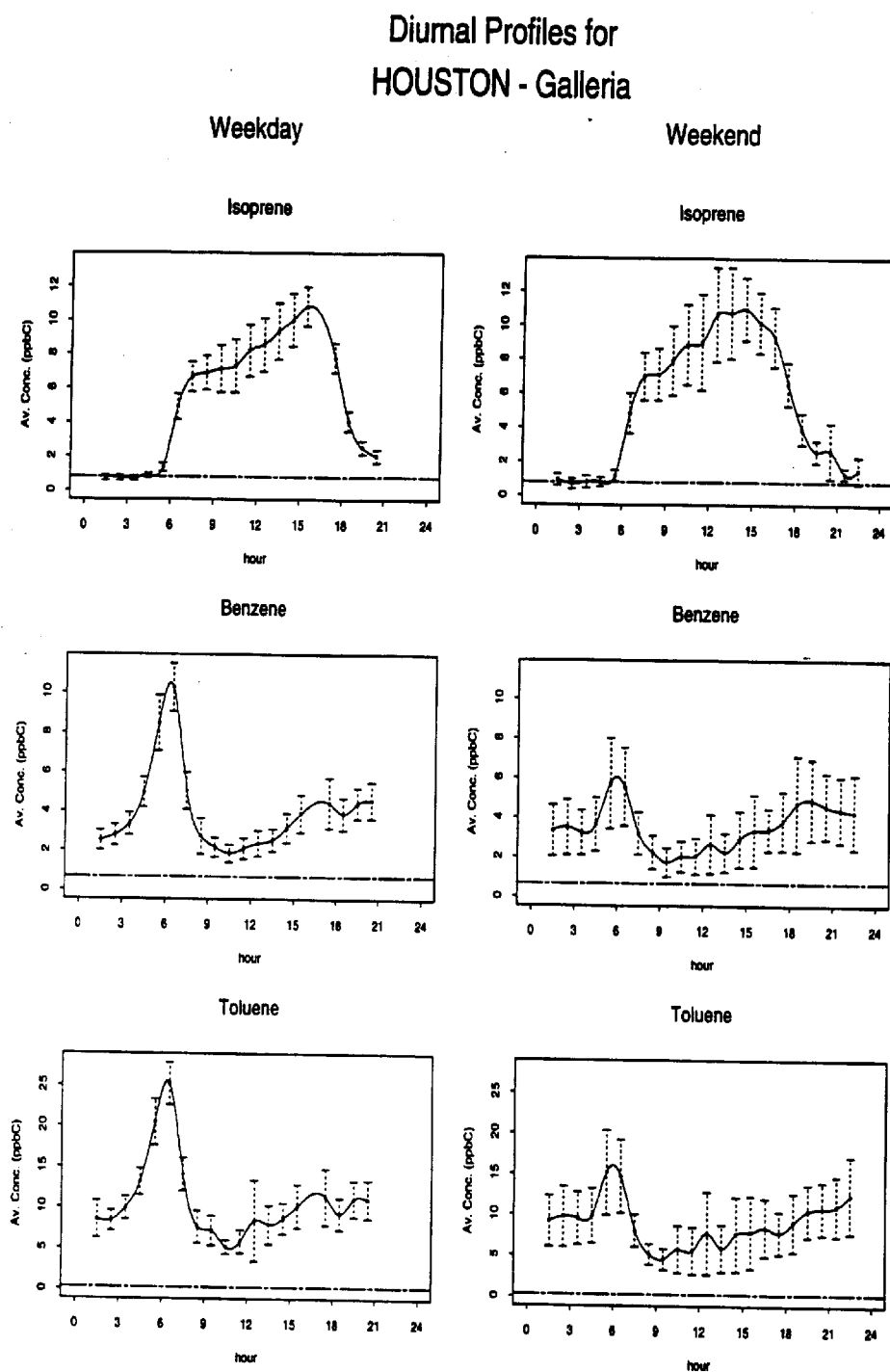
Figure 20-19.

BALTIMORE PAMS DATA 1993

DIURNAL BOX PLOTS--ORGANICS

Figure 21-20a.

Diurnal Profiles for
HOUSTON - Galleria

Weekday   Weekend

Acetylene   Acetylene

Ethylene   Ethylene

Ethane   Ethane

Figure 22-20b.

Diurnal Profiles for
HOUSTON - Galleria

Weekday   Weekend

Isoprene   Isoprene

Benzene   Benzene

Toluene   Toluene

Figure 23-20c.

Diurnal Profiles for
HOUSTON - Galleria

Weekday

Weekend

Olefins

Olefins

Xylene

Xylene

Figure 24-21a.

Diurnal Profiles for
HOUSTON - Galleria

Weekday   Weekend

Acetylene   Acetylene

Ethylene   Ethylene

Ethane   Ethane

Figure 25-21b.

Diurnal Profiles for
HOUSTON - Galleria

Weekday

Weekend

Isoprene

Benzene

Toluene

Isoprene

Benzene

Toluene

Figure 26-21c.

Diurnal Profiles for
HOUSTON - Galleria

Weekday

Weekend

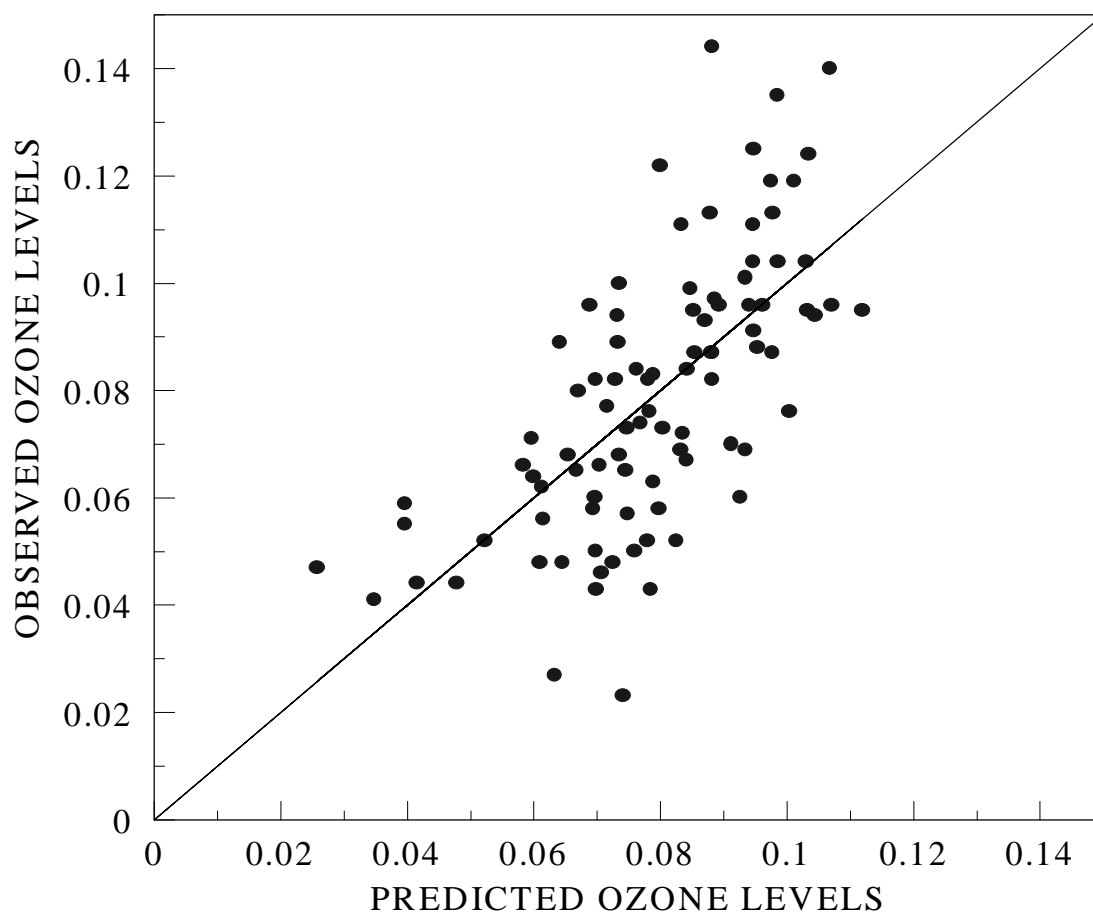#### Olefins

Olefins

Xylene

Xylene

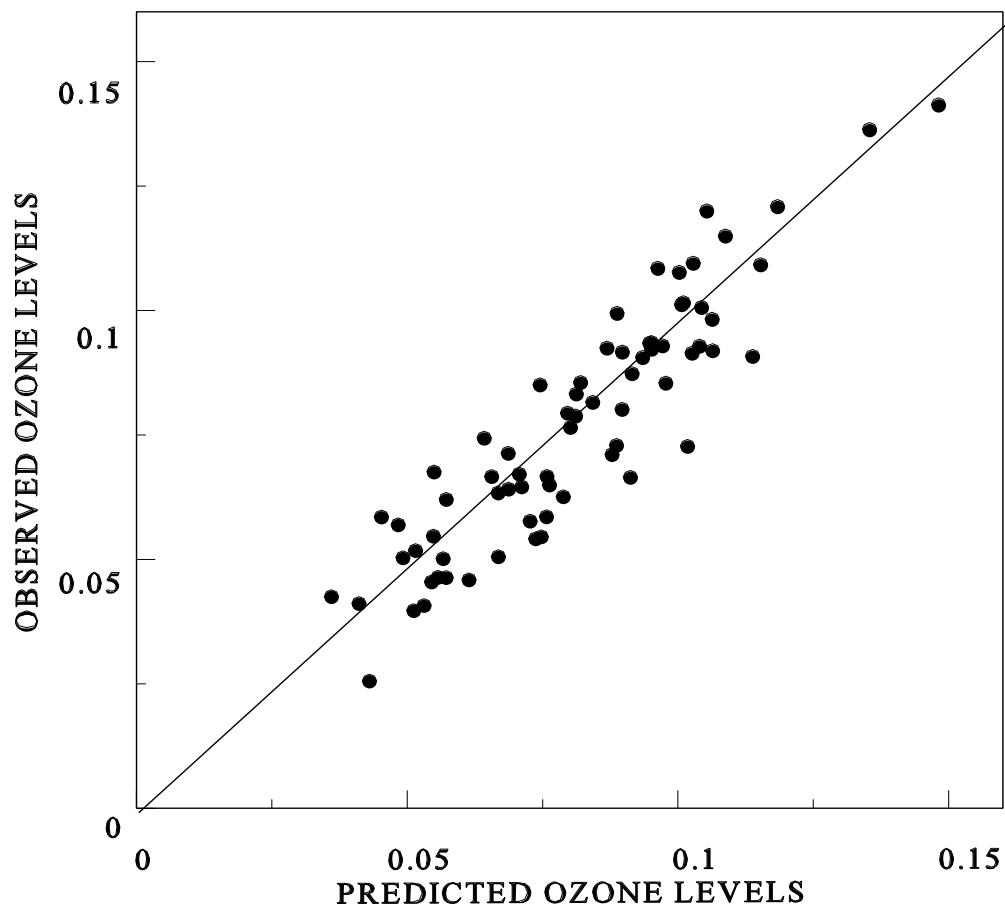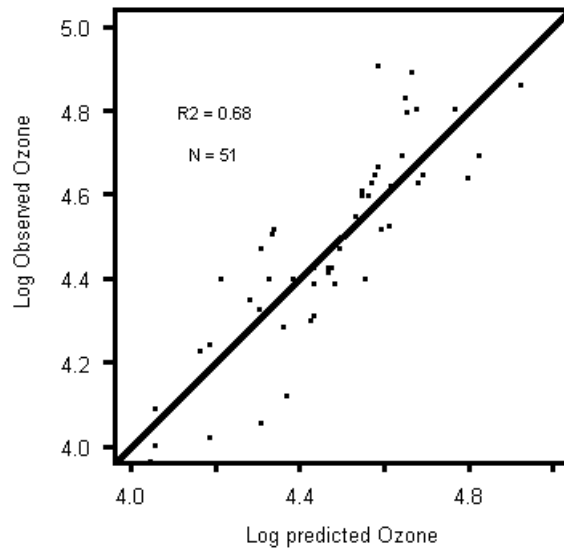**Figure 1-22.** Model 1- Meteorological Data Only; 1994 PAMS Data; Phialdelphia, PA; R2=0.45, n=92.

Figure 1-23.   Model 2 - Meteorological and Ozone Precursor Data; 1994 PAMS Data; Philadelphia, PA; R2=0.84, n=71.

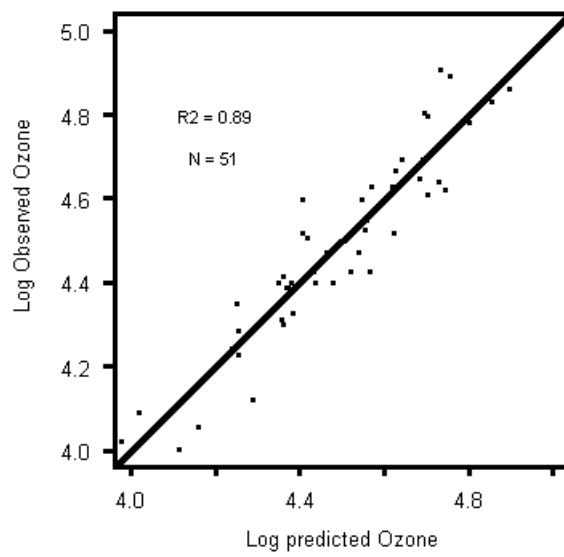**Figure 29-24. Philadelphia PAMS -Neural Network Size=1 Met, VOC and NO$_x$**

Figure 30-25. Philadelphia Neural Network Size=3--Met, VOC, NOX

Table 1-1. Most Abundant Anthropogenic VOCs in Selected Measurement Campaigns

	Northeast '94	5 City '84	Los Angeles '87
1.	Isopentane	Isopentane	Propane
2.	Toluene	n-Butane	Isopentane
3.	Propane	Toluene	Toluene
4.	Ethane	n-Pentane	n-Butane
5.	n-Butane	m/p-Xylene	Ethane
6.	m/p-Xylene	Propane	m/p-Xylene
7.	n-Pentane	Isobutane	n-Pentane

1.